

RE-THINKING GRAMMAR:
THE IMPACT OF EMBEDDED GRAMMAR TEACHING ON STUDENTS' WRITING
AND STUDENTS' METALINGUISTIC UNDERSTANDING

Debra Myhill – University of Exeter; *Professorial Fellow University of Wollongong, Australia*

Susan Jones - University of Exeter

Helen Lines - University of Exeter

Annabel Watson - University of Exeter

Corresponding author: Debra Myhill

Word Count: 11,836 excluding abstract, references and figures

Abstract:

This paper reports on a national study, involving a mixed methods research design comprising a randomised controlled trial, text analysis, student and teacher interviews and lesson observations. It set out to investigate whether contextualised teaching of grammar, linked to the teaching of writing, would improve student outcomes in writing and in metalinguistic understanding. The RCT involved 744 students in 31 schools in the South West and the Midlands of England, and was a blind randomisation study. Classes were randomly allocated to either a comparison or intervention group, after the sample had been matched for teacher linguistic subject knowledge. The statistical data was complemented by three interviews per teacher and three interviews with a focus student in each class, plus three lesson observations in each class, giving a data set of 93 teacher interviews, 93 student interviews and 93 lesson observations. In addition, the final pieces of writing produced for each scheme of work were collected. The statistical results indicate a significant positive effect for the intervention, but they also indicate that this benefit was experienced more strongly by the more able writers in the sample. The regression modelling also indicates that teacher linguistic subject knowledge was a significant mediating factor in the success of the intervention. The qualitative data provides further evidence of the impact of teacher knowledge on how the intervention was implemented and on students' metalinguistic learning. It also reveals that teachers found the explicitness, the use of discussion, and the emphasis on playful experimentation to be the most salient features of the intervention. The study is significant in providing robust evidence for the first time of a positive benefit derived from the teaching of grammar, and signals the potential of a pedagogy for a writing which includes a theorised role for grammar.

1.0 INTRODUCTION

The debate about the place of grammar in the English curriculum has a long history, with research reports and professional arguments on the topic spanning over fifty years. Moreover, it is a debate which crosses national boundaries and is common to most Anglophone countries. Detailed reviews of international evidence for and against the benefits of teaching grammar, stimulated by a renewed emphasis on grammar in the National Curriculum for English and the National Literacy Strategy, have been conducted by Hudson (2004), Wyse (2004) and most recently, by the EPPI Review Group for English (2004). It is a debate which has not only been theoretical and pedagogical, but one in which the *'public have regularly and enthusiastically participated'* (Gordon 2005). At the same time, there are international concerns about children's standards of writing. In Australia, following the 1996 National School English Literacy Survey, the Minister for Schools acknowledged that too many children did not achieve *'a minimum acceptable standard in literacy'* (Masters and Forster 1997). In the US, the call for a writing revolution (NCW 2003) to address the problem of children who *'cannot write with the skill expected of them today'* (NCW 2003), has been followed by major policy change in *No Child Left Behind* (US Department of Education 2002). This has reintroduced grammar as part of the *'raising standards'* agenda.

2.0 THEORETICAL FRAMEWORK

In tandem with political and ideological constructions of the grammar debate is an academic discussion between linguists and educationalists on the value of explicit and systematic teaching of grammar. The discipline of linguistics is, of course, much broader than the topic of grammar, including amongst other things, phonetics, pragmatics, evolutionary linguistics, language acquisition and neurolinguistics. In general, educationalists focus only on grammar, the structure of the language, and in some cases, on language acquisition. Increasingly, when discussing grammar, linguists draw on the principles of contemporary linguistic theories, which are descriptive and socio-cultural in emphasis, or as Carter describes them, *'functionally oriented, related to the study of texts and responsive to social purposes.'* (Carter 1990:104). They contend that a better understanding of how language works in a variety of contexts supports learning in literacy. Denham and Lobeck (2005) draw on empirical studies to claim that, in the multicultural, linguistically diverse classrooms of the US, linguistic knowledge is a tool which can inform teachers' approaches to language study in the classroom. Hudson (2004) offers a theoretical argument that education needs linguistics, noting the distinction between traditional prescriptive grammar and the very different approaches of modern linguistics. But teachers and educationalists remain sceptical. The minutes of a meeting of the Linguistic Society of America with the National Council for the Teaching of English, a powerful body representing English teachers, *'to discuss how to better integrate linguistics into the English/Language Arts curriculum'* note that NCTE was *'not eager to step in as partners in such a project (initiated by linguists)'* (LSA 2006:1). Indeed, many educationalists not only lack enthusiasm for grammar teaching but see it as potentially detrimental to children's language development. Braddock et al's review (1963) of research into composition concluded that teaching grammar was harmful, and Elbow argued that *'nothing helps [their] writing so much as learning to ignore grammar'* (Elbow 1981:169). As is evident from the publication dates already cited, this has been a long-running debate which remains unresolved.

The nature of the evidence base

There are, however, many conceptual and methodological flaws in much of the extant research base used to provide evidence for this debate. Research is repeatedly used selectively to justify a pre-determined position or to support a particular stance. Both Hudson (2004) and Wyse (2004) use research evidence to support their opposing standpoints, drawing both on differing evidence for their claims and offering different interpretations of the same

study. Bateman and Zidonis (1966) note that although most research in this area produces inconclusive results, these are then almost always construed as negative results. Tomlinson (1994), critiquing the methodological validity of the research of Robinson (1959) and Harris (1962), noted that the conclusions from their studies were '*what many in the educational establishment wanted to hear*' (Tomlinson 1994:26). Indeed, Tomlinson claims that most research into the effectiveness of grammar teaching does not stand up to '*critical examination*' and many articles are '*simply polemical*' (Tomlinson 1994:20).

The EPPI systematic review of the impact of grammar teaching on writing (EPPI 2004) highlights many of the flaws in this field of enquiry. The research question which informs the study (*What is the effect of grammar teaching on the accuracy and quality of 5 to 16 year olds' written composition?*) takes an over-simplified view of causal relationships between grammar teaching and written composition. It ignores the multi-faceted nature of learning and the complex social, linguistic and cognitive relationships that shape learning about writing. In particular, it does not engage with some of the key factors related to teaching and learning which might have an effect on the results. Teacher beliefs about the value of grammar, their level of linguistic and pedagogic subject knowledge, and teacher effectiveness in the classroom are important variables which are not considered. The background to the study does note that research might be needed to consider '*the effect/impact on students' writing skills of teachers' grammatical knowledge*' (EPPI 2004:12) but no subsequent account is taken of this. Likewise, there is no adequate conceptualization of 'grammar teaching': although the background to the study begins to explore the changes in linguistic theories and the international political and pedagogical trends in grammar teaching, these are not used to reformulate the research question. Thus the review considers research on grammar teaching but ignores the considerable and significant differences between the teaching of grammar in the UK in the early twenty-first century, and the teaching of grammar in different countries, in different decades, and in different contexts. The '*clear conclusion*', that '*there is no high quality evidence to counter the prevailing belief that the teaching of the principles underlying and informing word order and 'syntax' has virtually no influence on the writing quality of 5-16 year olds*' (EPPI 2004:4) is actually predicated on just three studies rated of medium or high significance, two of which are at least twenty-five years' old (Elley et al, 1975; Bateman and Zidonis 1966) and none of which were conducted in the UK. One of the review team subsequently reflected that '*our published reviews begged a lot of questions*' (Locke 2005:3)

Grammar and writing

Although there is a considerable number of international studies purportedly investigating the impact of grammar teaching on writing, there is almost none in which the grammar is taught in the context of writing lessons with a view to developing children's writing. In many of the studies (for example, Robinson 1959, Elley et al 1975, 1979; Bateman and Zidonis 1966) isolated grammar lessons are taught and the writing used to determine impact is produced in a different context. Fogel and Ehri's (2000) study is perhaps unique in taking as its starting point an identified writing problem, the tendency of some ethnic minority children to use non-standard Black English Vernacular (BEV) in their writing. The study set out to '*examine how to structure dialect instruction so that it is effective in teaching SE forms to students who use BEV in their writing*' (Fogel and Ehri 2000:215) and found a significant improvement in avoidance of BEV in the group who were given both strategies and guided support. They argue that their results demonstrated that the approach used had '*clarified for students the link between features in their own nonstandard writing and features in SE*' (2000:231). The Fogel and Ehri study moves the field forward by beginning to look at the pedagogical conditions which support or hinder the transfer of grammatical

knowledge into written outputs. It signals the importance of taking greater account of the subject knowledge of the teacher and undertaking a more fine-grained analysis of pupils' linguistic learning.

Teachers' linguistic subject knowledge

Concerns that teachers' linguistic knowledge is insufficient are neither new nor restricted to the UK. Gurrey observed that teachers lacked *'a thorough training in pure grammar'* (1962:14) and more recent concerns about the level of linguistic knowledge of English teachers have been expressed by Hudson (2004) in the UK, and by Koln and Hancock (2005) in the US, and Gordon (2005) notes teachers in New Zealand recognized *'their own, inadequate linguistic knowledge'* (Gordon 2005:50). These latter observations are, however, made by linguists, not teachers. A QCA (1998) survey of teachers in the period immediately following the introduction of the National Literacy Strategy indicated considerable lack of confidence in linguistic knowledge, particularly with sentence grammar, and uncertainty about implicit and explicit knowledge. The report concluded that there was a *'significant gap... in teachers' knowledge and confidence in sentence grammar and this has implications for... the teaching of language and style in texts and pupils' own writing'* (QCA 1998:35). From a pedagogical perspective, linguistic subject knowledge is more than the ability to use appropriate terminology, as it also involves the ability to explain grammatical concepts clearly and know when to draw attention to them. Andrews suggests that it is *'likely to be the case that a teacher with a rich knowledge of grammatical constructions and a more general awareness of the forms and varieties of the language will be in a better position to help young writers'* (Andrews 2005:75), and Gordon (2005) found that teachers who developed more secure linguistic knowledge were able to see beyond superficial error in children's writing to evidence of growing syntactic maturity. Previously, for these teachers *'the "writing virtues" of their pupils often went unseen and unacknowledged because of their own lack of knowledge about language'* (Gordon 2005:63). In contrast, weak linguistic knowledge can lead to an over-emphasis upon identification of grammar structures without fully acknowledging the conceptual or cognitive implications (Myhill 2003) of that teaching. Equally, it can lead to sterile teaching, divorced from the realities of language in use: Applebee, for example, notes two studies in the US which showed that topic sentences and paragraph patterns taught in school bear little resemblance to those found in 'real' prose (Applebee 2000).

The role of metalanguage

Central to the issue of linguistic subject knowledge and the debate about the role of grammar in developing writing is the question of the value of grammatical terminology and access to this metalanguage. Cognitive psychologists have repeatedly signalled the importance of metacognition (Hayes and Flower 1980; Martlew 1983; Kellogg 1994; Wallace and Hayes 1992; Butterfield et al 1996) in the writing process, because writing is a process which demands self-monitoring (Kellogg 1994:17). Metacognition, put simply, is thinking about one's thinking: it is a conscious process, described by Flavell as *'one's knowledge concerning one's own cognitive processes or anything related to them, e.g., the learning-relevant properties of information or data. For example, I am engaging in metacognition if I notice that I am having more trouble learning A than B; if it strikes me that I should double check C before accepting it as fact'* (Flavell 1976:232). Metacognitive knowledge plays a role in every stage of the writing process: in moving planning from an over-emphasis on content to greater consideration of the strategic goals of the task (Hayes and Flower 1980); in supporting the development of *'a model of their audience, for reflecting on rhetorical and content probabilities'* (Kellogg 1994: 213); in the process of revision (Alamargot and Chanquoy 2001); and in developing self-regulation (Englert et al 1992). Bereiter and Scardamalia (1982) argue that the benefit of metacognition is that it renders *'normally covert processes overt'* and provides *'labels to make tacit knowledge more accessible'* (1982:57) and in summarizing the findings of their intervention study, Englert et al are insistent that *'the*

importance of the students' increased mastery over the language of the writing process cannot be over-emphasised' (1992:441): both are thus signalling the importance of a metalanguage, though not necessarily grammatical language. Metalinguistic knowledge is a subset of metacognitive knowledge, though there is surprisingly little empirical research in this aspect of writing. The most comprehensive consideration of metalinguistic knowledge is Gombert's (1992) model of metalinguistic awareness, designed to inform an understanding of oral development and how children learn to read. He proposes two levels of cognitive control of linguistic knowledge: *epilinguistic*, where linguistic processing is controlled automatically by linguistic organisations in the memory, and *metalinguistic*, when, the individual is in conscious control of linguistic decision-making. Gombert argues that there is a *'developmental hierarchy between epilinguistic control and metalinguistic awareness'* (2003). Van Lier (1998), however, contests the hierarchical assumptions of Gombert's proposition, questioning whether epilinguistic awareness is necessarily a precursor of metalinguistic awareness. Yet this principle of a cognitive shift from implicit to explicit knowledge is a prevalent one, including at policy level. QCA (1998) describe the learning trajectory as moving from implicit knowledge, derived from experience, to analysis, based on grammatical terminology, developing into understanding of function and effect, leading finally to explicit knowledge. Explicit knowledge is defined as *'knowledge that can identify and account for connections and distinctions between different examples of usage, enhance reading and improve writing'* (QCA 1998:20). Van Lier questions the value of metalinguistic knowledge *'measured in solitary demonstrations of knowledge'* and argues that being able to articulate metalinguistic knowledge is less important than being able to demonstrate it: the ability *'to control and manipulate the material at hand'* is more significant than the ability *'to describe a linguistic feature using grammatical terminology'* (1998:136). Van Lier's concern that metalinguistic knowledge is not transferred into linguistic performance is central to the issue of whether grammar supports writing development. Myhill and Jones (2007) found that secondary age writers were often able to articulate explicit choices made during text production, but were not always able to describe these in metalinguistic terms; equally, they found that some proficient writers had automated linguistic decision-making and no longer thought explicitly about metalinguistic choices. On the other hand, Carter (1990) maintains that the demise of formal grammar teaching and with it the absence of a metalanguage in the classroom has been disempowering, preventing learners from *'exercising the kind of conscious control and conscious choice over language which enables both to **see through** language in a systematic way and to use language more discriminatingly'* (Carter 1990:119).

This review of research into the relationship between grammar and writing indicates that as Andrews observed, *'there is still a dearth of evidence for the effective use of grammar teaching of any kind in the development of writing'* (Andrews 2005:74). It is evident that there remains a pressing need for robust large-scale research which seeks to establish valid causal relationships, but which also seeks to go beyond simple cause-effect paradigms to understand the complexity of the issue. In particular, writing research needs to adopt an inter-disciplinary framework, which is cognisant of linguistic, cognitive and socio-cultural perspectives, in order to reflect with validity the complexity of classrooms as teaching and learning contexts. It is important to be mindful of the cognitive demands of writing production, and the challenge all writers face of keeping in mind *'the conceptual message together with their rhetorical objectives and at the same time appeal to linguistic knowledge to express their ideas correctly and appropriately'* (van Gelderen 2005:215). However, it is equally important to foreground the linguists' perspective *'that terminology and rules are pointless if your mind hasn't grasped the concepts behind the terminology'* (Keith 1997:12). In addition, the cognitive and linguistic challenges of writing need to be bounded by an acknowledgement that writing is *'material social practice in which meaning is actively made, rather than passively relayed or effortlessly produced'* (Micciche 2004: 719). The study reported here sought to operate

within such an inter-disciplinary framework and to answer the research question: *What impact does contextualised grammar teaching have upon pupils' writing and pupils' metalinguistic understanding?*

3.0 METHODOLOGY

Teaching is a complex, multi-faceted and situated endeavour which resists simplistic causal explanations between pedagogical activity and learning outcome; equally, writing is perhaps the most complex activity learners undertake, drawing on cognitive, social and linguistic resources. Accordingly, this study adopted a mixed-method approach located within an inter-disciplinary conceptual framework, combining a cluster randomised controlled trial (RCT) with multiple regression analysis and a complementary qualitative study.

The EPPI review of the effect of grammar teaching (EPPI: 2004) concluded by calling for '*a conclusive, large scale and well-designed randomised controlled trial*' (EPPI: 2004:49) into the impact of grammar teaching on writing. We would argue that no such randomised controlled trial (RCT) could be conclusive because of the complexity of both the empirical question and the educational context. Instead, this study adopted a mixed methods approach to investigate specifically the complex causal relationships between pedagogical support for teaching grammar, teacher subject knowledge and improvement in writing. One intervention can be realised in multiple outcomes for learners and one intervention strategy can be multiply interpreted and mediated in the classroom by different teachers. Indeed, as Shadish, Cook and Campbell argue, causal relationships are rarely deterministic: '*to different degrees, all causal relationships are context dependent*' (Shadish et al 2002:5). Therefore, to complement the statistical data derived from the experimental study, and to provide in-depth understanding of the theoretical, pedagogical and contextual implications of the statistical data, the experimental component was embedded within a qualitative design. This mixed method approach is important for RCTs conducted in educational contexts: indeed, Moore, Graham and Diamond (2003) argue that '*to undertake a trial of an educational or social intervention without an embedded qualitative process evaluation would be to treat the intervention as a black box, with no information on how it worked, how it could be improved, or what the crucial components of the intervention were.*' Likewise, Shadish, Cook and Campbell (2002:71) recommend '*the addition of qualitative methodologies to experiments*' to provide better interpretation and avoid errors in applying research outcomes to practice.

The collection of data fell broadly into three categories: the baseline data required before the intervention began; the impact data which measured the effect of the intervention; and the qualitative data which provided interpretative and contextual information. The impact data comprised the pre and post-test writing scores, including both the total scores, and the scores on the sub-components of the writing test. The qualitative data comprised lesson observations of the sample classes, teacher interviews, writing conversations with the sub-sample of pupils, and the writing produced in the teaching episodes.

3.1 The randomised controlled trial

In educational settings, conventional RCTs following the medical model are rarely appropriate as frequently interventions are at the level of the school or the class, rather than the individual. Instead, the cluster trial where '*randomisation is conducted at the level of the group or cluster*' (Moore, Graham and Diamond 2003:680) is increasingly viewed as a more valid and robust design. This study utilized a cluster RCT where the independent variable was the pedagogical support materials and the dependent variable was the quality of writing. The

intervention group received detailed pedagogical support materials and were trained in their use, whilst the comparison group received only an outline scheme of work with no pedagogical support.

3.1.1 The sample

In this study, as a cluster trial, randomisation occurred at the level of the group rather than the individual (Murray, 1998; Donner & Klar, 2000). Only one class was used from each school to avoid any cross-over effects from one group to another. Neither the teacher nor the class knew there was a distinction being made between an intervention and a comparison group. The teachers and students knew that the focus of the research was writing, but they were unaware of the specific grammar focus. The student sample was 32 year 8 mixed ability classes in comprehensive schools with between 24 and 30 students in each class.

The process of recruiting the schools made use of Local Authority school directories in the South West and the West Midlands. In order to create as representative a sample as possible, an initial process de-selected the small number of schools which were single-sex, selective, or atypical in age range (eg 14-18 age range). This was because these schools could distort the sample in terms of gender, attainment or socio-economic status. This left a sample which comprised mixed, comprehensive schools with an age range of 11-18. Using a random number generator, each school was given a number, creating a numbered list of the schools. Schools were then contacted in numerical order to invite them to participate in the project, and if there was a non-response or a decline to participate then the next school on the list was approached. Random sampling continued until at least 16 schools from each area had been recruited. 19 schools in the South-West and 28 schools in the West Midlands were approached to secure the sample of 16 in each area.

Because the study set out to investigate the impact of teachers' linguistic subject knowledge (LSK) upon the intervention, the sample was stratified first by teacher linguistic subject knowledge in order to avoid accidental bias in either the intervention or comparison group. The LSK of the teachers was established through a specifically designed test of grammar knowledge in a questionnaire establishing baseline information about the teacher's professional background and subject knowledge. It included questions regarding professional and academic background; questions focusing on teachers' perceptions of their own knowledge of English Literature and their beliefs about the use of literature to support the teaching of writing; questions focusing on teachers' grammar knowledge and their beliefs about the use of grammar to support the teaching of writing. The score for each teacher on the linguistic subject knowledge test was added to the database. The teacher scores were ranked and then in turn allocated to either group A or group B, thus ensuring that each group was broadly matched for LSK. Finally, the two groups were randomly allocated to the comparison or intervention group.

Having established the participating schools, teachers and classes, a comprehensive set of baseline data was collated (see Table 1). This drew heavily on national test data, collected in all schools, and comparable across the sample. At student level, data on attainment in English was gathered. Students in England sit national externally-marked tests at age 11 (Key Stage 2 tests) and are required to report a National Curriculum level for English at age 14 (Key Stage 3). At Key Stage 2, the test score is available as an overall National Curriculum level for English, plus a separate level and raw score for writing. The student data also recorded whether a child was in receipt of Free School Meals or whether English was an Additional Language for them. At teacher level, the baseline data was collected through the questionnaire described above. At school level, the data was drawn from the most recent OFSTED inspection report, which provides comparable contextual information about each school inspected.

Insert Table 1 here

3.1.2 *The intervention*

The intervention comprised detailed teaching schemes of work in which grammar was embedded where a meaningful connection could be made between the grammar point and writing. Both the comparison and intervention groups taught the same writing genre over a three week period once a term, and addressed the same writing learning objectives from the Framework for English, part of the English government's National Strategies for raising educational attainment (see Table 2). Both groups were given the same written outcomes for each genre studied: the opening of a story; a written speech; and a portfolio of three specified types of poem. A medium term plan was provided for each group, which outlined the time frame, learning objectives, assessed outcomes, accompanied by a range of relevant stimulus resources. Thus, for both the intervention and comparison groups, the learning focus, the period of study, the learning objectives, and the assessed written outcomes were the same.

Insert Table 2 here

The detailed teaching schemes for the intervention group were designed by the project team, and explicitly sought to introduce grammatical constructions and terminology at a point in the teaching sequence which was relevant to the genre being studied; for example, exploring how the use of first or third person can position the narrator differently or looking at how expanded noun phrases can build description in poetry (further information about the teaching strategies can be found in Myhill et al 2011). The teaching focus was on effects and constructing meanings, not on the grammatical terminology: building on the concept of writing as design (Sharples 1999, Myhill 2010), the goal was to open up what we have called '*a repertoire of possibilities*', rather than to suggest correct or formulaic ways of writing. A set of pedagogical principles informed the design of the teaching schemes:

- ❑ The grammatical metalanguage is used but it is always explained through examples and patterns
- ❑ Links are always made between the feature introduced and how it might enhance the writing being tackled
- ❑ The use of 'imitation': offering model patterns for students to play with and then use in their own writing
- ❑ The inclusion of activities which encouraging talking about language and effects
- ❑ The use of authentic examples from authentic texts
- ❑ The use of activities which support students in making choices and being designers of writing
- ❑ The encouragement of language play, experimentation and games

3.1.3 *Pre and post test writing tasks*

The impact of the teaching on student writing was determined by a pre and post-test sample of writing. Both the pre and post-test writing sample were a first person narrative, drawing on personal experience, and written under controlled conditions. The test design and marking was led by Cambridge Assessment, who were responsible for setting and marking the national Writing Test at Key Stage 3 until 2006. In order to avoid any possible bias created by the precise choice of writing task, the topic was selected to avoid known gender preferences in writing and to avoid any need for having had a particular experience. To ensure that there was no task bias, a cross-over design was adopted where half the sample completed task 1 as the pre-test and task 2 as the post-test, while the other half of the sample reversed the order in which these tests were taken. Both sample sets were independently

marked by Cambridge Assessment. For each set of scripts, Cambridge Assessment provided a first marker's set of marks, a second marker's set and a 'resolution mark', adjudicated by a third senior marker if the first two marks were very different. They indicated which set of marks should be used for the purposes of the study, but allowed us to see any variability in the marking behind those chosen figures. The marking was based on the national Key Stage 3 mark scheme format, the final mark being made up of three components: sentence structure and punctuation; text structure and organization; and composition and effect. Cambridge Assessment devised the training materials for marking; undertook the administration to select and train a marking team; delivered a training day for each marking round; and ensured the usual standardisation checks during the marking. The markers did not know from which treatment group the writing had derived.

3.1.4 Attrition and fidelity

Fidelity is a problematic concept in a naturalistic educational setting such as this, as identical implementation of the intervention teaching materials is neither possible nor desirable. Teachers were not asked to follow the lesson plans rigidly; they were allowed to adapt materials to suit the needs of their students, but were also asked to remain as close as possible to the materials. All 32 participating schools remained in the project throughout the year long period of the study, but it was decided to exclude one school's data from the final analysis because of low fidelity to the study. For example, the teacher taught lessons which were not focused upon writing, and also regularly had other teachers teaching her lessons. Therefore, the final sample used for analysis comprised 31 teachers in 31 schools.

The original student sample was $n=900$, but after removing one class on the grounds of low fidelity, and removing those students who were not present for both the pre-test and the post-test, the final sample was $n=744$, representing an attrition rate of 16.2%.

3.2 The qualitative study

The qualitative data is comprised of four data sets: classroom observations and follow-up interviews once per teacher per scheme of work; writing conversations once per scheme of work with a focus child from each class; and the written outcomes from the teaching of each genre. The full qualitative data set includes 93 lesson observations; 93 teacher interviews; 93 writing conversations, plus the writing samples of narrative fiction, argument and poetry from each class.

For the lesson observations, a schedule was designed to capture a record of how the teachers taught the three writing genres and how students responded. The schedule recorded the sequence of activities and the grammatical, literary and linguistic terminology used by the teacher, and prompted for comments on teacher interaction, student responses and observers reflections on the use of contextualised grammar teaching.

The teacher interviews were conducted using a semi-structured interview schedule which explored their pedagogical decision-making in the lesson observed, and their reflections on the lesson and students' learning. For the intervention group only, this included their pedagogic evaluations of the schemes of work. The specific questions about the lesson observed were not pre-planned but drew on the lesson observation, for example, probing why the teacher gave a particular example, or the teacher's reflection on a particular student's response. In addition, each interview sought to generate teachers' beliefs about writing, through prompts designed to stimulate open-ended discussion. The final teacher interview, following the poetry scheme of work, was an

extended interview which directly explored teachers' beliefs about the value of grammar teaching: this was not directly addressed in the earlier interviews to avoid highlighting grammar as the focus of the study, although many of the teachers did refer to grammar without prompting in the earlier interviews.

The writing conversations with a student were conducted following the lesson observation. In each of the sample classes, one student was selected as a focus student for these conversations; the overall sample of focus students was stratified by gender in order to avoid any data distortions which might be influenced by gendered attitudes towards writing. The writing conversations were shaped by a semi-structured interview schedule and by stimulus prompts. The first section of the interview was principally to facilitate engagement by asking broad questions about writing and their own perceptions of learning about writing, but it did include a question which probed their perspectives as learners on the lesson observed. As with the teacher interviews, the precise questions asked at this point were framed by what had been observed in the lesson. The second section of the interview explored students' metalinguistic understanding of their own and others' writing. This section of the interview used stimulus prompts to lead the discussion. For each genre, a sample of writing in that genre written by students of the same age was selected, not as a model of excellence but as a starting point for discussion. The interview also used the focus student's own writing produced during the teaching as a stimulus for discussion.

The qualitative data enabled the study to be sensitive to the complexities of classroom learning, and to provide a more nuanced picture of the way the intervention was realized in practice. It will both *'inform future development of the intervention, and also to contribute to theory and understanding of the relation between context, mechanism and outcome'* (Moore, Diamond and Graham 2003).

3.3 Ethical considerations

Ethical considerations have been informed by the institutional Research Ethics policy, the BERA Revised Ethical Guidelines for Educational Research (2004) and the ESRC Research Ethics Framework. The original proposal underwent an institutional ethical review and was awarded a Certificate of Approval.

However, the blind randomisation design created a particular ethical problem as it was not possible to tell participants which treatment group they were in, or the precise focus of the study. All participants were informed that the study was researching writing, but not that it was investigating the impact of grammar teaching on writing. Thus the 'informed' consent was partially compromised. In order to address this, all participants were informed at the outset that all research results and a full outline of the conduct of the research would be communicated to them at the end of the study.

A further ethical problem specific to this study was the need to *'minimise the effects of designs that advantage ... one group of participants over another'* (BERA 2004:8): half the teachers received pedagogical support which could give their classes an educational advantage/disadvantage. This was addressed in two ways. Firstly, the period of trialling the support materials attempted to remove any threat of negative impacts upon student learning during the main study, and both groups were given matching learning objectives, stimulus resources and written outcomes. Secondly, at the end of the project the outcomes of the study were disseminated directly to participating schools through a teacher conference so that any beneficial impacts could be adopted more widely.

4.0 THE OUTCOMES OF THE STUDY

4.1 The impact of grammar teaching on students' writing

The outcome variable considered in the analysis was the difference between the post and the pre-test percentage marks for each of the 744 participants in the study. The results indicate that both intervention and comparison groups improved over the time period of the study, with the overall mean value of the improvement between post- and pre-test scores being 9.24%. However, for the intervention group (n=412), the mean outcome was 11.52%, which contrasts with a mean outcome of 6.41% for the comparison group (n=332). So overall, and ignoring the effect of any other covariates, a simple two-sample t-test suggests a highly significant ($p < 0.001$) positive difference of 5.11 percentage marks for the intervention in terms of improvement in writing attainment. In layperson's terms, those students in the intervention group improved their attainment in the post-test writing test significantly more than the comparison group. This represents the first robust statistical evidence for a beneficial impact of the teaching of grammar in students' writing attainment.

However, the simple overall comparison above needs to be refined to allow for significant effects of school, teacher and student explanatory covariates, so as to obtain a more nuanced understanding of the effect of the intervention. To achieve this, multiple regression analysis was applied to the individual student level data (n=744). The difference in percentage marks between the post and pre writing test scores for each individual was again used as the outcome variable. The explanatory covariates were firstly, whether the individual was in the intervention or comparison group, and then the baseline measures (see Table 1) at student, teacher and school level, where school and teacher covariate values were attributed to each student according to his or her school or class membership. Interaction terms between these covariates were also considered in the modelling. As the covariates included both categorical measures (eg Free School Meals) and continuous measures (eg linguistic subject knowledge), the multiple regression modelling could be described more specifically as a multi-way analysis of covariance.

A range of model selection procedures were then used to determine which variables and interactions were statistically significant in impacting the response from amongst the intervention indicator and the other student, teacher and school level covariates. The model selection was by guided both by theoretical considerations and size of estimated co-efficients, as well as purely statistical procedures such as 'stepwise regression' techniques and formal statistical tests of differences in explanatory power between competing models. Final model selection was then checked using full residual analyses and standard diagnostic procedures. The results from the model selection indicated that the most appropriate reduced model involved the explanatory variables/factors or interactions listed in Table 1. All terms in this model are very highly statistically significant (at a 0.1% level) with the exception of the teacher LSK score which is significant only at the 5% level. None of the factors, interactions or variables not reported in Table 1 were statistically significant at the 5% level and, more importantly, a formal F-tests shows no significance difference in model fit between the reduced model in Table 1 and the full model containing all of the original variables and interactions. Overall, the reduced model explains only 9% of the raw variation in the pre- and post test differences ($R^2 = 0.09$, adjusted $R^2 = 0.08$), but low overall explanatory value is not atypical in educational studies and taken overall this model is very highly significant ($F_{7,36}$ statistic=10.42, p -value<.00001)

Insert Table 3 here

4.1.1 Relationships between the intervention and student writing attainment

Table 3 shows that the intervention had a more marked positive effect on able writers. It also shows that, over the period of the intervention, able writers in the comparison group made less progress in writing than less able writers, whereas able writers in the intervention group made significant progress. The implications of these results are firstly, that able writers receiving conventional teaching of writing may be stalling in their progress at age 14, and secondly, that the attention to the relationship between grammar and writing may be particularly appropriate to their learning needs. The intervention may have been pitched too much towards able writers: it drew on understanding of students' linguistic development from a previous study (Myhill 2008; 2009). Further research would be necessary to establish whether using the same pedagogic strategy of embedded grammar teaching but addressing different aspects of writing more relevant to lower attaining writers' needs would be more successful. It is also possible that able writers had higher levels of cognitive understanding of the concepts being taught, and that they were more effective in transferring the learning into their own writing.

4.1.2 *Relationships between the intervention and teachers' linguistic subject knowledge and experience*

The research design set out to investigate not only whether contextualized grammar teaching might support learning about writing and improve writing attainment but also whether teachers' confidence with grammatical knowledge, or their length of teaching experience might influence the intervention in any way. The statistical data indicates that both these factors were significant mediating factors. Students in intervention classes with the most and the least experienced teachers benefited less than those teachers with 5-10 years experience. There is no supporting evidence or illumination in the qualitative data to explain this result, but it is possible that teachers with the most experience were more resistant to altering firmly embedded practices, and that more recently qualified teachers followed the intervention materials in a more technicist manner, focusing on the activities rather than the learning. Further research would be required to draw more robust conclusions.

In terms of linguistic subject knowledge, those students in intervention classes with teachers with higher subject knowledge benefitted more than those with teachers who had lower subject knowledge. This is a less surprising finding, because of the very obvious pedagogical relationship between linguistic subject knowledge and the nature of the intervention. The teaching materials provided some support for linguistic knowledge through the accompanying resources, but nevertheless required confident mastery of grammar, particularly in making meaningful connections for writers between a linguistic construction and a piece of writing, and in being able to cope with children's questions. The qualitative data provides considerable further evidence of this relationship between the intervention and the teacher's linguistic knowledge, and will be discussed in more detail later in this paper.

Overall, then, the statistical analysis indicates a positive impact of the use of contextualized grammar teaching on student writing. To date, this is the first large-scale study which has found this result. However, the statistical data also indicates differential effects which are important both for theory, policy and practice. Firstly, the intervention was more beneficial to able writers than weaker writers; secondly, the linguistic subject knowledge of the teacher was a significant influencing factor; and thirdly, that length of teaching experience influenced learner outcomes. This data allows us to construct scenarios which describe the different factors which shape the way the intervention plays out in practice. So, for example, a student who is an able writer in a class with a teacher of 5-10 years experience and good linguistic subject knowledge is likely to benefit more from this pedagogical approach than a student who is a weaker writer in a class with a recently qualified teacher whose linguistic subject knowledge is limited. These differences are important when considering the implications of the study for policy or

practice as they underline the need to take more account of contextual factors beyond an apparent causal relationship between the pedagogical approach and writing improvement.

4.2 The impact of pedagogical support materials on the teaching of grammar

Given the overall positive results of the intervention, but also the evidence that other factors, such as linguistic subject knowledge, were influential it is important to understand how the teachers used the materials and their views of them. In the teacher interviews, teachers reflected on how they had used the intervention materials; in addition, the interviewer used the lesson observation notes to invite discussion of particular aspects of the lesson relevant to the focus on grammar and writing. Nvivo was used to analyse the codes inductively: the coding process derived two over-arching categories from the data, each with its own themes. These were firstly, comments which addressed teachers' views and use of the materials; and secondly, comments which reflected teachers' perspectives of student response to the materials (see appendix). The themes will be italicized in the text below for ease of identification.

4.2.1 Teacher reflections on their own practice using the materials

Two themes, *Resources/Activities* and *Timing*, were more general comments on the materials. Many teachers had found the pace of the lessons too fast, and noted timing issues. That notwithstanding, the comments about the resources and activities were highly positive: teachers liked the schemes of work and even found them '*brilliant to share with the other teachers*'. These comments tell us little specifically about the intervention, other than confirming professional approval. However, this approval appeared to assist *Fidelity* to the teaching intervention. More than half the teachers maintained they had followed the materials closely because '*the scheme of work is really, really good*' or because students were succeeding in the lessons and '*really learning*'. The theme, *Adaptation*, gave an insight into the ways adjustments were made; although we wanted teachers to follow the schemes closely, it was recognised that teachers had to be free to adapt to meet the needs of their students. Many adaptations were pragmatic, often to timing, or exchanging one resource for another similar one. But it is relevant to the research focus that some of the adaptations were to avoid or circumvent the grammar focus because the teacher was not confident with the material or because the teacher was concerned '*it would have just completely have gone over their head*'.

This concern about the grammar element of the intervention materials was also reflected in the theme, *Linguistic Subject Knowledge*, in which approximately half of the teachers expressed their anxiety about how well they could handle grammar explanations or student questions. One teacher reflected that '*I didn't find it easy because I'm struggling to get my head round understanding some of these things myself and I think sometimes it shows and I think sometimes the kids know, and sometimes they throw out answers that I'm not quite sure about and I think my unsureness comes across.*' At the same time, teachers felt the materials had given them *Confidence* in various aspects of writing pedagogy such as '*the way that I teach imaginative writing*', but significantly, no teacher directly argued that the teaching materials had developed their confidence with grammar.

One of the pedagogical principles behind the design of the teaching materials was being explicit in highlighting grammar features where they related to the writing being taught. Teachers were not informed about the pedagogical principles until the end of the project so it is noteworthy that, after comments on the practical aspects of the materials, the theme of *Explicitness* was the most frequently articulated. Teachers spoke about how the materials had been more explicit about linguistic features than their own normal practice, using words such as

explicitly, focused and specific repeatedly. Whilst teachers felt they frequently taught students about vocabulary and word choice, the emphasis on syntactic features in the materials was noted as less usual. Explicit use of grammatical terminology was also noted, with several teachers reflecting that they had never taught determiners and modal verbs. For one teacher, there was a recognition that the use of terminology supported students' articulation of their understanding about writing, when previously *'they hadn't perhaps been given the terminology to be able to express it very clearly'*. Many of the comments in this theme identified this explicitness as a distinguishing feature of the materials, different from standard practice, and in some cases participation in the project had changed their viewpoints: *'I've changed my mind about this one, it is crucial to teach children explicitly how to write well'*.

4.2.2 Teacher reflections on student responses to the materials

A substantial set of responses referred to high levels of *Student Engagement* with the materials, and there is probably a correspondence between this and the teachers' approval of the materials. There may well, therefore, be halo effects here which would not be sustained in the longer term. There appeared to be an element of *Surprise* at students' responses; as one teacher admitted, *'students have engaged at a much higher level than I was expecting first off'*. Some of this surprise related to students being able to cope with activities their teachers had not expected them to manage. This links to another strong theme in the teacher reflections, that some things were *Too Difficult* for the students. Aspects perceived as difficult always related to the grammar focus: one teacher's students *'struggled with the noun phrases'*, whilst another teacher's students found the idea of viewpoint *'quite difficult to grasp and they kept saying 'Is it first person or third person?''* Conversely, a high number of comments reflected the development of student *Understanding*, and the development of explicit knowledge about grammar and writing: *'they obviously got the whole thing about connectives as to which ones work best when and why and what have you'*. Significantly, this understanding was purposeful knowledge applied to writing; as one teacher noted, *'They've far more understanding of what it is they're doing it and why they're doing it'*.

Four of the themes have direct links with the pedagogical principles underpinning the intervention. Teachers enjoyed the *Discussion* opportunities provided in the schemes and reported *'brilliant discussions'* where *'students were willing to risk opinions about language more'*. They recognised that the discussion helped to create transferable knowledge, *Making Connections* for learners between the grammar and the writing, and generating greater learner independence where students were *'making a link without me'*. This chimes with the aims of the teaching materials which sought to avoid generating formulaic knowledge about writing and which endeavoured to open up students to a repertoire of possibilities to draw on in writing. One strategy to support this was the use of activities which encouraged language play. Teachers felt that this *Experimentation* had created space for valuable play, and particularly required that they handed over *'control a little bit and allow them to just experiment'*. Teachers also believed that students were *Making Informed Decisions* about their writing, and that they had understood the importance of *'making choices and making decisions about your reading and your writing'*.

Finally, at least half the teachers believed the teaching materials had supported *Writing Improvement*. It is important to be cautious in claims made on the basis of this as the research project was obviously addressing writing improvement and again there may be halo effects here. For some teachers there was a perception of a direct relationship between the intervention and the improvement: *'I do feel that all this work we've been doing on working with words, working with sentence structures, working with how you work with these different word classes to enhance your writing, I'm sure that's fed into it because it was beautiful work'*.

In conclusion, then, the evidence from the teacher interviews suggests two levels of impact of the intervention teaching materials on the teaching of grammar. Firstly, the teachers found the materials high-quality, stimulating strong student engagement. These are qualities which could apply to any teaching materials and are not specific to the grammar intervention. The second level of impact is specific to grammar teaching. The teachers' comments suggest that some of the pedagogical principles underpinning the planning had significant impact on their teaching. The explicitness of the attention to linguistic features built into the teaching materials and the creation of multiple opportunities for discussion are the most prominent of these, but the emphasis on playfulness and experimentation, and on developing repertoires of possibility through making connections and encouraging informed decision-making are also important. Conversely, in addition to expressing concern about their linguistic knowledge, some teachers demonstrated their anxiety in practice by avoiding some of the grammar points in the lessons, and their strong sense that the materials might be too difficult for students may reflect their own difficulty as much as student difficulty.

4.3 The impact of the intervention on students' metalinguistic understanding

The student interviews included the writing conversations, where the student was invited to comment on a prompt piece of writing of the same genre under study, and to reflect on their own writing so far. The coding of students' metalinguistic understanding drew on Gombert's taxonomy (1992) of metalinguistic understanding as the basic framework for analysis. Gombert identified six categories of metalinguistic understanding (metaphonological; metalexical; metasemantic; metasyntactic; metapragmatic; and metatextual), which will be summarized here as word, sentence and text level understanding. For a fuller account of this aspect of the study, see Myhill (2011a). Students in both the intervention and comparison groups were given the same prompts and the same opportunities to reveal metalinguistic understanding: however, students in the intervention group made more comments defined as evidence of metalinguistic understanding, and made more elaborated comments. This implies that one element of the success of the intervention was in developing metalinguistic understanding about writing. But their confidence in articulating metalinguistic understanding was not equal across word, sentence and text level domains, as will be outlined below.

4.3.1 Word level metalinguistic understanding

Without doubt, the greatest confidence in discussing linguistic choices in writing was at word level. The students were very aware of the impact and significance of word choices, understanding that words helped '*create a picture in your mind*' or contributed to the creation of a particular viewpoint, showing '*what the person in the story is actually seeing*'. They understood the potential emotiveness of word choices, of '*words that are going to sink into the heart*' and how the choice of appropriate vocabulary '*makes you feel something that the writer wants you to feel*'. When talking about his own piece of argument writing, one writer observed that he had chosen the words '*downhearted, inhumane, defenceless, neglected*' to describe a character in order '*to make people feel sorry for him*'.

The prompt text for argument included a mix of formal and informal vocabulary: for example, the writer used '*elderly*' rather than '*old*', which is a more formal option, and conversely used the verb '*carted*', rather than an alternative such as '*taken*'. Many students recognised that word choices were important in increasing the formality of text, and several commented on the choice of the word '*older generation*' rather than '*older people*', a

choice which one students described as ‘not *posh but pretty civilised*’. The word ‘*disgraceful*’ stimulated many comments, which often combined a recognition that the word was emotive and a more formal choice. As one student observed: ‘*Disgraceful is a really strong word to use ... instead of saying ‘Oh it’s horrible’, because that probably wouldn’t grab your attention if you just said ‘it’s horrible’ because it’s quite a common word to use, lots of people use it.*’

4.3.2 Sentence level metalinguistic understanding

Gombert’s (1992) definition of metasyntactic understanding is that it represents conscious reasoning about syntax and deliberate control over accurate use of grammar. One difficulty with Gombert’s taxonomy is that it relates to oral development and is particularly focused, therefore, on young learners and how they develop grammatical accuracy in speech. In our study, only the weakest writers presented grammatical inaccuracies in their writing, and then only rarely. More relevant to this study was these students’ ability to understand and discuss how syntactic choices helped writers to meet their rhetorical goals. The teaching schemes drew explicit attention to aspects of syntactic variety as part of a repertoire for shaping text: the possibilities of short sentences for impact, varying sentence lengths to create textual rhythm, and about altering the syntactic structure of sentences to shift focus. Overall, however, it is this aspect of metalinguistic understanding which seemed to be least secure in this sample.

Greatest confidence was shown in discussing the use of short sentences for particular effects and in recognising the value of variety in sentence length within a text. The way a short sentence can draw attention to itself as a kind of prosodic exclamation seemed to be understood: students talked of short sentences which ‘*grab attention*’ and ‘*make[s] you think about it*’. However, whilst the students have grasped that variety might be a good writer’s choice, they are less assured in articulating any reasons linked with prosody or textual rhythm. Indeed, in some cases, it was thought to be more to do with supporting reader understanding than creating an effect: *If there’s a really long sentence you’re probably thinking ,oh well I can’t remember what the beginning of the sentence was’.*

Understanding the possibilities of syntactic variety in altering how a sentence is read was much less well understood, despite this being one focus of the teaching. Some writers were able to make fairly precise comments about syntactic choices, as this writer evidences: ‘*I’ve put the connective at the start like I said before and I wouldn’t normally do that, normally it would be ‘The pig will have nowhere to sleep’ or something like that, but I’ve put ‘Despite the fact that the pig will have nowhere to sleep’.* However, this writer was unable to explain why this choice had been made and for what purpose. It is possible that this writer, and indeed many of the other students in the study, do not yet have sufficient explicit syntactic knowledge to articulate their emergent metasyntactic understandings.

4.3.3 Text level metalinguistic understanding

Gombert (1992) argued that metatextual understanding, a more global awareness of coherence and cohesion, develops later than metalexical and metasyntactic understanding. There were fewer comments coded as metatextual in this study, but the teaching focus was directed more to word and sentence level aspects of writing, which may account for this. In the poetry unit, we had anticipated there might be more opportunities for metatextual discussion, because the poem as a text is very visible as a whole and often deliberately organized to exploit textual possibilities. However, students seemed to struggle to talk about a poem as a text, focusing more on layout than structure, for example, the writer who observed that the poem ‘*couldn’t be a story because it’s too short and it couldn’t just be a passage of writing, because each sentence starts the same way.*’

There was more evidence in the narrative and argument units of students having awareness of the text as a whole. In the narrative scheme, one writer articulated a clear text level decision that *'It's going to be told in first person by one of the tramps called Toby'* and another explained her use of tense to manage different time points: *'it will be past at the beginning from when he's telling the story of his past but towards the main bit it will be present, so it will go from past to present.'* Students were also confident explaining plot development and narrative structure, including understanding that narrative can be non-chronological and *'you don't necessarily have to put the opening at the beginning, you can put it in the middle'*. In the argument scheme, students were developing awareness of how the argument was structured across the text: *'They've put the important points at the top and important points at the bottom and in the middle they've kind of tried to persuade you it's bad. They've put a summary of the whole kind of speech in a couple of sentences at the end which is good, kind of refreshes your memory'*.

In summary, the data not only shows that the intervention had a positive impact upon developing metalinguistic understanding but also signals that metalinguistic learning about writing is socially-constructed, and in this case heavily shaped by what teachers value in writing. Many of the comments, both those demonstrating strong metalinguistic understanding and those showing more limited understanding, frequently mirrored closely what teachers had said in the lessons. In particular, because some of the teachers did not have sufficient linguistic subject knowledge to handle metalinguistic discussion confidently, students' understanding was correspondingly limited. Many of the teachers were more comfortable teaching about word choices than syntactic variety, which, as will be outlined below, reflects relatively less assurance with syntax than with word grammar.

4.4 The impact of teacher linguistic subject knowledge on the teaching of grammar

The statistical data, as reported earlier, indicated that teacher linguistic subject knowledge was a mediating factor in influencing student outcomes. The teachers' scores on the Linguistic Subject Knowledge test were very evenly spread across the cohort (see Table 5) with a mean result of 60% and a fairly high Standard Deviation of 15.8%. Only one of the teachers in this sample had a degree which included a linguistics element, and this teacher, not surprisingly, scored 86%. Given that older teachers may have been taught grammar as part of their own education, the results were analysed to see if there was any relationship between years of teaching experience and linguistic knowledge, but there is no strong evidence that this is so. The top four results do include three teachers with more than 28 years experience, but equally the second lowest result is from a teacher with 23 years experience.

Insert Table 4 here

The lesson observations in particular, but also the teacher and student interviews, provided explanatory evidence of how linguistic subject knowledge impacted upon the teaching. Overall, there were three ways in which a lack of confidence in grammar was realized in the classroom or in reflections on practice:

- Making meaningless comments about grammar
- The use of semantic definitions
- Syntactic confusions

4.4.1 Meaningless grammar

Teacher comments to students during observed lessons sometimes included advice to writers which either made no sense at all, or was insufficiently elaborated or explained to be meaningful for students. One teacher told her students that *'if you use verbs, adverbs or nouns, you will be able to write a very powerful description'*, which is not helpful – it would be hard to write at all without using verbs, adverbs, or nouns and, moreover, it is perfectly possible to write weak and ineffective descriptions using verbs, adverbs and nouns. Another set of less helpful comments related to the idea of sentence variety, which was a teaching focus of the schemes of work. Teachers regularly advocated the use of variety: *'variety is important'*; *'make sure you have sentence variety'*. However, there was rarely any explanation of why this variety was beneficial, and implying that variety, of whatever quality, was a good thing. In contrast, one teacher with good linguistic subject knowledge, gave a more precise and meaningful reason for using varied sentence lengths which made a link between the linguistic feature and how it might impact upon the writing. In the argument scheme of work, she was discussing how students could use contrast in sentence length in different ways: *'in a long sentence you can detail the cruelty and a short sentence you can refer to sudden death for impact'*.

Another tendency was to promote the use of a particular linguistic feature 'for effect':

- *think about where you put your punctuation for effect*
- *use sentences for effect*
- *vary vocabulary for effect*
- *short sentence used for effect*

To an extent, these reflect the teaching materials which repeatedly encouraged discussion about the effects of grammar features, but many teachers lacked the applied linguistic knowledge which allowed them to move beyond the phrase 'for effect' to a more text or context specific discussion of the possible effects created.

4.4.2 *The use of semantic definitions*

Another pattern of response which links to lack of confidence in handling linguistic terminology was the strong tendency to give students semantic definitions for word classes, rather than linguistically precise descriptions. So verbs were regularly defined as *'doing'* words, thus leading to student difficulties when they encountered verbs which do not appear to involve any action (for example, *are; will; wonder; consider*), but especially when they encountered words which were not verbs but which implied an action (as in *'I am deeply opposed to hunting'* where many children identify 'hunting' as the doing word because of its implied action). Frequently, the semantic definitions offered were partial, such as adverbials are *'size words'*, or a noun is *'the name of an object'*. Adjectives were regularly defined as *'describing words'*, without acknowledgment of the descriptive power of lexical verbs, nouns and adverbs. The tendency to use semantic definitions is influenced by common practice over many years in both primary and secondary classrooms, practice which is often endorsed in commercial materials. Occasionally, however, the semantic definition was completely unique as was the case with the teacher who described an adverb as an *'action plus word'*.

These semantic definitions led to some very confused discussions with students because they held on very tightly to the semantic definition and applied it with absolute logic. In one lesson observed, the teacher had earlier suggested that a noun was something you can touch, which resulted in the following exchange:

Teacher: *What are the rules for whether it is a noun or a verb or something else?*

Student: *It's if you can touch it.*
 Student: *Can you touch it?*
 Student: *Can you go to it?*
 Teacher: *Can you touch hockey? But hockey is a noun.*
 Student: *You play hockey so it must be a doing word.*
 Teacher: *Is her a noun?*
 Student: *Yes, you can touch her.*
 Teacher: *Can safe be a noun?*
 Student: *Safe is a feeling not a thing.*

Sometimes the students recognised the flaw in these semantic definitions. In one student interview, one girl suggested she could improve her writing by adding more adjectives to strengthen the description, but then added *'If you think about it, all words are adjectives because they're all describing things. A noun is describing'*.

4.4.3 Syntactic confusion

The aspect of the Linguistic Subject Knowledge test in which teachers scored least well was the set of questions on clauses and syntax, suggesting this is an area of particular challenge. In the interviews, many teachers articulated a specific anxiety about clauses. In the lessons observed, teachers often chose to focus on sentence variety in terms of sentence length, as this is very easy to handle, and requires no grammatical explanations. Many ungrammatical variations on sentence types were developed, with the grammatical distinctions of simple, compound and complex being extended with concepts such as *'more simple'*; *'very complex'*; *'a normal sentence'*; and a *'more than average sentence'*. Linked to the issue of using semantic definitions for word classes, described above, at a syntactic level, the concept of grammatical simplicity was confused with semantic simplicity, so simple sentences were short sentences and complex sentences were long sentences. This led to one student offering this explanation in an interview:

Student: ***'There was a cat'** is a simple sentence; a complex sentence is like, **'There was a slim, something, something ginger cat.'***

In contrast, some teachers demonstrated more confident management of discussion of syntactic features which linked them very explicitly to the way they were working in the specific piece of writing under focus. In the narrative fiction scheme of work, one teacher responded to a student's draft with the feedback that *'I like the way you've kept some short sentences in to build the tension'*, and another drew attention to the way adverbials can create a sense of place and setting: *'We've got a real sense of the environment with adverbials in there'*. On other occasions, teachers deftly drew attention to the subtleties of making changes to the standard Subject-Verb order of a sentence:

- *Look at this and the way it's been changed. Sometimes you can change the structure of a sentence to make it more interesting.*
- *Look what's happened by changing the word order. As a writer you can withhold information and build a sense of expectation.*

These data reveal how basic problems with declarative knowledge of linguistic metalanguage, particularly syntactic knowledge, and knowledge about the mobility of word classes in English generate very real problems in working constructively with grammar in the context of writing. Where teachers' own subject knowledge was limited, there were frequently corresponding applied pedagogical problems in providing adequate definitions or explanations of linguistic terms. The reliance on semantic explanations, rather than functional explanations, often confused students, and teachers sometimes found it hard to handle student questions or conceptual problems with confidence. Significantly for the focus of this study, limitations in linguistic subject knowledge meant that some teachers struggled to make meaningful links for students between a linguistic feature and its effect or purpose in a specific text. Conversely, where teachers had greater command of the linguistic subject knowledge, they were better able to make purposeful connections between grammar and writing, and were more confident managing discussion about effects and possibilities. Teachers with confidence in linguistic subject knowledge helped writers shape text creatively; teachers who lacked confidence provided formulaic recipes for success.

5.0 IMPLICATIONS FOR THEORY, POLICY AND PRACTICE.

The study represents the first large-scale study in any country of the benefits or otherwise of teaching grammar within a purposeful context in writing. It stands in contrast to previous studies which were either small-scale (Fogel and Ehri 2000; Bateman and Zidonis 1966) or which investigated whether discrete grammar instruction improved writing outcomes (Elley et al 1975; 1979), and is the only study of its kind conducted in England. It is also significant in combining complementary qualitative and quantitative methods, which have provided not only statistical evidence of the impact of the intervention on writing outcomes, but also a richly nuanced understanding of how the intervention worked. The intervention materials were theoretically informed by the principles of writing as design (Kress 1994; Kress and van Leeuwen 1996; Sharples 1999; Cope and Kalantzis 2000), in which creators of text make design choices from an available repertoire of possibilities. The use of grammar as a design tool (Myhill 2010; 2011b) elaborates thinking about writing as design beyond the visual and the multimodal to include specific linguistic decision-making. Alongside this, therefore, the intervention also sought to develop students' metalinguistic understanding through explicit instruction and through opportunities for discussion. The strong positive effect of the intervention signals for the first time the potentiality of grammar as an enabling element in writing development and evidences a clearly theorised role for grammar in writing pedagogy. However, the result that the intervention benefited able writers most is also important. Further studies should usefully investigate whether teaching materials designed around the identified linguistic learning needs of less able writers would have a beneficial effect; and whether the use of materials which incorporate explicitness and discussion of linguistic focuses but without the use of the metalanguage might be more supportive to less able writers.

The study also highlights the importance of teachers' linguistic subject knowledge in mediating metalinguistic knowledge in the writing classroom. The teacher interviews and lesson observations indicate that teachers need to possess confidence in declarative knowledge of grammar, particularly syntactic knowledge, if they are to be able to handle students' questions and misunderstandings effectively. This would include an ability to define and explain metalinguistic terminology appropriately. This is an aspect of grammar teaching which has been systematically overlooked at both policy level and in research; there is neither understanding nor agreement about how best to explain grammatical terminology. At the same time, however, the study makes it clear that declarative knowledge alone is insufficient. As Parr (2009) argues, one facet of pedagogical subject knowledge for the teaching of writing is *'the ability to articulate and make accessible to developing writers that which is implicit and often at a level below conscious'* (2009:147); on the one hand this implicit knowledge may be about writers

and the process of writing, but on the other hand, it is about texts and how texts work. Teachers need to be able to apply their linguistic subject knowledge to published texts and to children's own writing, identifying significant linguistic features and being able to make connections for writers between a feature and its impact on a text or reader.

At policy level, the study suggests judicious caution about too simplistic an advocacy of or legislation for a specific pedagogical practice. Whilst there is now robust evidence from the data in favour of the use of grammar in an embedded way within the teaching of writing, the study certainly does not suggest that this would be of universal benefit. Rather it emphasizes the complex inter-relatedness of many factors in the realization of educational benefit; particularly in terms of learners' needs, teachers' attitudes and experience, and teachers' subject knowledge. Policy development needs to take these interacting factors into account and, in particular, consider how to develop professional pedagogical ownership of policy in ways which foster principled adaptation to meet learners' needs and interests.

Finally, we think it is important to acknowledge the limitations of the study, and especially the RCT design. Whilst there are benefits in the large-scale quantitative data, the RCT may be too focused on causal relationships between an intervention and student outcomes, paying insufficient attention to other factors. In seeking to generalize, it always has the potential to miss the particular. The emphasis on principles such as intervention fidelity, blind randomization and bias can serve to exclude the very variables which are most significant should the intervention be generalized into professional practice. Further studies which build on this study might consider the use of a series of smaller-scale, but nonetheless statistically robust, interventions which combine the testing of the impact of specific interventions with full involvement of the teachers in understanding the pedagogical principles underlying the intervention. In particular, we would recommend that teachers design and develop the teaching materials for any intervention themselves, with guidance from the research team, thus taking ownership of the pedagogical principles which inform the study.

REFERENCES

- Alamargot, D., and L. Chanquoy. 2001 *Through the Models of Writing*. Kluwer Academic Publishers
- Andrews, R. 2005 Knowledge about the Teaching of Sentence Grammar: The state of play. *English Teaching: Practice and Critique*. 4 (3) 69-76
- Andrews, R., C. Torgerson, S. Beverton, A. Freeman, T. Locke, G. Low, A. Robinson, and D. Zhu. 2006 The effect of grammar teaching on writing development. *British Educational Research Journal* 32 (1) 39-55
- Applebee, A. 2000 Alternative Models of Writing Development. In *Perspectives on Writing: Research, Theory and Practice*, ed. R. Indrisano and J. Squire. Newark: IRA pp90-111
- Bateman, D. R., and F. J. Zidonis. 1966 *The Effect of a Study of Transformational Grammar on the Writing of ninth and tenth graders*. Champaign, Illinois: National Council of Teachers of English
- Bereiter, C., and M. Scardamalia. 1982 From conversation to composition: the role of instruction in a developmental process. In *Advances in Instructional Psychology Vol 2*, ed. R. Glaser. Hillsdale, NJ: Lawrence Erlbaum Associates pp1-64
- Braddock, R., R. Lloyd-Jones., and L. Schoer. 1963 *Research in Written Composition*. Urbana: National Council of Teachers of English.
- Burstein, L. 1980 The Analysis of Multilevel Data in Educational Research and Evaluation. *Review of Research in Education*. 8: 158-233.
- Butterfield, E.C., D.J. Hacker, and L. R. Albertson. 1996 Environmental, cognitive and metacognitive influences on text revision. *Educational Psychology Review*. 8 (3) 239-297
- Carter, R. (ed) 1990 *Knowledge about Language*. London: Hodder and Stoughton
- Cohen, J. 1988 *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cope, B. and M. Kalantzis. 2000 *Multiliteracies: Literacy learning and the design of social futures*. London: Routledge.
- Denham, K. and A. Lobeck. 2005 *Language in the Schools: integrating linguistics into K12 teaching*. Hillsdale, NJ: Lawrence Erlbaum Associates
- Donner, A. & N. Klar. 2000 *Design and Analysis of Cluster Randomization Trials in Health Research* London: Arnold.
- Elbow, P. 1981 *Writing with Power: Techniques for Mastering the Writing Process*. New York: Oxford
- Elley, W. B., I.H. Barham., H. Lamb, and M. Wylie. 1975 The Role of Grammar in a Secondary School Curriculum. *New Zealand Council for Educational Studies* 10 pp26-41 .
- Elley, W. B., I.H. Barham., H. Lamb, and M. Wylie. 1979 *The Role of Grammar in a Secondary School Curriculum: Educational Research Series No 60*. Wellington: New Zealand Council for Educational Research.
- Englert, C.S., T.E. Raphael, and L.M. Anderson. 1992 Socially Mediated Instruction: Improving Students' Knowledge and Talk about Writing. *The Elementary School Journal* 29 (4) 411-449
- EPPI Review Group for English. 2004 *The Effect of Grammar Teaching (Syntax) in English on 5 to 16 year olds' Accuracy and Quality in Written Composition*. London: EPPI-Centre
- Fogel, H. and L.C. Ehri. 2000 Teaching elementary students who speak black English vernacular to write in Standard English: effects of dialect transformation practice. *Contemporary Educational Psychology*. 25, 212-235
- Gombert, J. E. 1992 *Metalinguistic Development*. Hemel Hempstead, UK: Harvester-Wheatsheaf.

- Gombert, J. E. 2003 Implicit and Explicit Learning to Read: Implication as for Subtypes of Dyslexia *Current Psychology Letters: Behaviour, Brain and Cognition* 10 (1) Special Issue on Language Disorders and Reading Acquisition. <http://cpl.revues.org/document202.html> [accessed Mar 2006]
- Gordon, E. 2005 Grammar in New Zealand Schools: Two Case Studies. *English Teaching: Practice and Critique*. 4 (3) 48-68
- Gurrey, P. 1962 *Teaching English Grammar*. London: Longmans
- Harris, R.J. 1962 *An experimental enquiry into the functions of and value of formal grammar in the teaching of English*. Unpublished PhD thesis, London University
- Hayes, J.R., and L.S. Flower. 1980 Identifying the Organisation of Writing Processes. In *Cognitive Processes in Writing*, eds L.W. Gregg and E. R. Steinberg. New Jersey: Lawrence Erlbaum Associates, Hillsdale. pp3-30
- Hudson, R. 2004 Why Education needs Linguistics. *Journal of Linguistics* 40 (1) 105-130
- Keith, G. 1997 Teach Yourself English Grammar. *The English and Media Magazine*. No 36 Summer 1997
- Kellogg, R. T. 1994 *The Psychology of Writing*. Oxford: Oxford University Press
- Koln, M. and C. Hancock. 2005 The Story of English Grammar in US Schools. *English Teaching: Practice and Critique*. 4 (3) 11-31
- Kress, G., and T. van Leeuwen. 1996 *Reading Images The Grammar of Visual Design*. London: Routledge
- Kress, G. 1994 *Learning to Write*. London: Routledge
- Locke, T. 2005 Grammar Wars – Beyond a Truce. *English Teaching: Practice and Critique*. 4 (3) 1-10
- LSA. 2006 *Minutes of the Meeting of the Committee on Language in the School Curriculum*. LSA: Albuquerque http://www.lsadc.org/info/pdf_files/2006LiSCminutes.pdf [Accessed July 2006]
- Martlew, M. 1983 Problems and Difficulties: cognitive and communicative aspects of writing. In *The Psychology of written Language: Developmental and Educational Perspectives*, ed. M. Martlew. Chichester: John Wiley
- Masters, G. and M. Forster. 1997 *Literacy Standards in Australia*. Melbourne: Australian Council for Educational Research
- Micciche, L. 2004 Making a case for rhetorical grammar. *College Composition and Communication*. 55(4), 716-737.
- Moore, L., A. Graham and I. Diamond. 2003 On the Feasibility of Conducting Randomized Controlled Trials in Education. *British Education Research Journal*. 29 (5) 673-689
- Murray, D. 1998 *Design and Analysis of Group-Randomized Trials*. Oxford: Oxford University Press.
- Myhill, D.A. 2003 Principled Understanding? Teaching the Active and Passive Voice. *Language and Education*. 17 (5) 355-370
- Myhill, D. A. and S. Jones. 2007 More than just Error Correction: Children's Reflections on their Revision Processes. *Written Communication*. 24 (4) 323-343
- Myhill, D.A. 2010 Ways of Knowing: Grammar as a Tool for Developing Writing. In *Beyond the grammar wars: A resource for teachers and students on developing language knowledge in the English/literacy classroom*, ed. T. Locke. London: Routledge
- Myhill, D.A. 2011a 'The Ordeal of Deliberate Choice': Metalinguistic Development in Secondary Writers. In *Past, present, and future contributions of cognitive writing research to cognitive psychology*. ed. V. Berninger. Psychology Press/Taylor Francis Group. pp247-274
- Myhill, D.A. 2011b Grammar for Designers: How Grammar Supports the Development of Writing. In *Insight and Impact: Applied Linguistics and the Primary School*, eds. S. Ellis, E. McCartney, and J. Bourne. Cambridge: Cambridge University Press. pp81-92

- Myhill, D.A. Lines, H.E and Watson, A. (2011) Making Meaning with Grammar: a Repertoire of Possibilities *mETaphor* 2:1-10
- National Commission on Writing in America's Schools and Colleges. 2003 *The neglected "R": The need for a writing revolution*. New York: College Board.
- Parr, J. 2009 Building professional knowledge to teach writing: Teacher pedagogical content knowledge and student achievement in writing. Presented at the 45th UKLA Annual Conference, Literacy Today and Tomorrow. University of Greenwich, 10–12 July 2009.
- QCA. 1998 *The Grammar Papers: Perspectives on the Teaching of Grammar in the National Curriculum*. London, QCA
- Raudenbush, S.W. 1988 Educational Applications of Hierarchical Linear Models: A Review. *Journal of Educational Statistics*. 13 (2) 85-116.
- Robinson, N. (1959) The relation between knowledge of English Grammar and ability in English compositions. Unpublished MEd thesis, University of Manchester
- Shadish, W.R., T. Cook, and D. Campbell. 2002 *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin
- Snijders, T.A.B., & R. J. Bosker. 1993 Standard errors and sample sizes in two-level research. *Journal of Educational Statistics*. 18 (3) 237-260.
- Snijders, T.A.B. 2005 Sample Size in Multilevel Linear Models. *Encyclopedia of Statistics in Behavioral Science* 3 pp 1570-1573
- Tomlinson, D. 1994 Errors in the Research into the Effectiveness of Grammar Teaching. *English in Education*. 28 (1) 20-26
- United States Department of Education. 2002 *No Child Left Behind (NCLB) Act of 2001*. Washington, D.C.: US GPO.
- Van Gelderen, A. and R. Oostdam. 2005 Effects of fluency training on the application of linguistic operations in writing. *L1 Studies in Language and Literature*. 5 (2) 215-240
- Van Lier, L. 1998 The Relationship between Consciousness, Interaction and Language Learning. *Language Awareness*. 7 (2 &3) 128-145
- Wallace, D. L. and J. R. Hayes. 1992 Redefining Revision for Freshmen. In *Reading empirical research studies: the Rhetoric of Research*, eds. J.R Hayes and R.E Young, Hillsdale, New Jersey: Lawrence Erlbaum pp 352-369
- Wyse, D. 2004 Grammar. For writing? A Critical Review of Empirical Evidence. *British Journal of Educational Studies*. 49 (4) 411-427

Appendix: Summary of Coding Themes

Teacher-focused themes	Definition	Number of Responses
Explicitness	Comments claiming the schemes of work were more explicit about grammar or teaching writing than in their normal practice	41
Resources/activities	Comments about the quality of the resources and activities	59
Fidelity	Comments which seem to suggest the teacher has stuck to the schemes of work	10
Adaptation	Comments which refer to things the teacher changed or dropped from the Schemes of work	57
Subject knowledge concern	Comments which refer to teacher uncertainty about linguistic subject knowledge	21
Confidence	Comments which refer to how the schemes of work gave them confidence	7
Timing	Comments which refer to problems with timing or amount of material to be covered	34
Student-focused themes	Definition	Number of Responses
Surprise	Comments in which teachers express surprise at some aspect of students' work	19
Making connections	Comments which refer to how the schemes of work or students are making connections between grammar and writing	26
Making informed decisions	Comments which refer to students making conscious choices in writing	14
Experimentation	Comments which claim the schemes of work encouraged students to experiment or play with language	31
Discussion	Comments which refer to the schemes of work provoking student discussion	41
Writing improvement	Comments about ways in which the writing had improved or not	11
Student engagement	Comments which refer to activities or resources which engaged the students	43
Understanding	Comments which refer to students' understanding	45
Too difficult	Comments which refer to things not taught because of a perception they were too difficult for the students	58

	Data collected
Student	Key Stage 2 English level Key Stage 2 Writing level Key stage 2 Writing raw score Predicted Key Stage 3 English level Free School Meals or not English as an Additional Language or not
Teacher	Years of service Initial undergraduate degree subject Linguistic subject knowledge score
School	Free School Meals (average; above average; below average) Ethnic diversity (average; above average; below average) % of students achieving 5 GCSEs or more, including Maths and English % of students with Special Educational Needs Contextual Value Added Measure Most recent OFSTED inspection grade OFSTED Section 10 English result

Table 1 – A summary of the baseline data

Time frame	Writing genre	Learning Objectives for Writing from the Framework for English Year 8
Autumn Term	Narrative Fiction	Developing viewpoint, voice and ideas Varying sentences and punctuation for clarity and effect Improving vocabulary for precision and effect Developing varied linguistic and literary techniques Using grammar accurately and appropriately
Spring Term	Persuasive Writing	Developing viewpoint, voice and ideas Varying sentences and punctuation for clarity and effect Improving vocabulary for precision and effect Developing varied linguistic and literary devices Structuring, organising and presenting texts in a variety of forms on paper and on screen Using grammar accurately and appropriately
Summer Term	Writing Poetry	Generating ideas, planning and drafting Varying sentences and punctuation for clarity and effect Improving vocabulary for precision and impact Developing varied linguistics and literary techniques

Table 2: Learning Objectives addressed in the schemes of work

Model coefficient	Estimate	Std. error	t-value	p-value
Intercept	-6.1727	(3.5285)	(-1.739)	(0.723)
Student level				
Above average writing level	-7.4006	(1.9037)	(-3.887)	(<0.001)
Intervention group and above average writing level	8.1246	(1.3863)	(5.860)	(<0.001)
Teacher level				
Teacher Linguistic Subject Knowledge (LSK) score	0.5817	(0.2760)	(2.108)	(0.031)
Teacher has 5-10 years experience	5.9832	(1.3091)	(4.570)	(<0.001)
School Level				
'Satisfactory' result in last School Inspection	7.0165	(1.9258)	(3.643)	(<0.001)
'Good or excellent' result in last School Inspection	8.3394	(2.0151)	(4.138)	(<0.001)
% SEN in School	0.6411	(0.1610)	(3.983)	(<0.001)

Table 3: coefficient estimates (standard errors, t-values and p-values) for student, teacher and school level covariates in final selected model

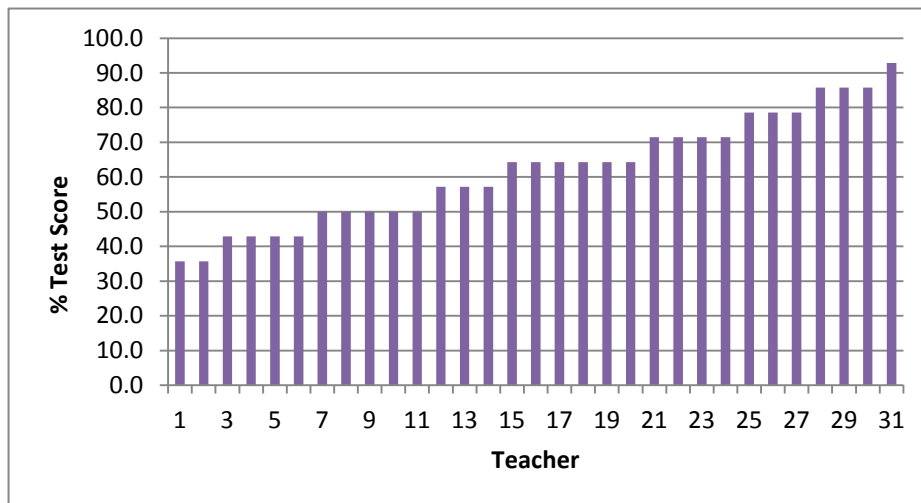


Table 4: teachers' linguistic subject knowledge scores